

WHAT IS CLAIMED IS:

1. A method for generating a structural interaction fingerprint (SIFt) in the form of an information string which comprises a plurality of information blocks wherein each information block comprises a plurality of information units, the method comprising:
 - selecting a plurality of selected positions on a target molecule, wherein each selected position corresponds to an information block in the information string, the target molecule forming a complex with a ligand;
 - selecting a plurality of interaction types and calculating a value that is indicative of a characteristic of each interaction type at each selected position of the target molecule;
 - assigning the value to a corresponding information unit, the information unit indicating a characteristic of the interaction type at the corresponding selected position;
 - joining the information units of each selected position together to form corresponding information blocks; and
 - joining the information blocks together to generate a SIFt.
2. The method of claim 1, wherein the target molecule is a protein or a peptide.
3. The method of claim 1, wherein the target molecule is a nucleic acid.
4. The method of claim 1, wherein the ligand is a small molecule, a peptide, a protein or a nucleic acid.
5. The method of claim 1, wherein the value that is assigned to an information unit is a binary value, which indicates the presence or absence of a particular interaction type at the corresponding selected position.
6. The method of claim 1, wherein the value that is assigned to an information unit is a numeric value selected from a scale of numbers, wherein the numeric value indicates the magnitude of a particular interaction type at the corresponding selected position.
7. The method of claim 1, wherein the selected positions are obtained from a three-dimensional structure of a binary complex formed between the target molecule and the ligand.

8. The method of claim 7, wherein the three-dimensional structure is derived from an experimental method or a prediction method.
9. The method of claim 2, wherein each selected position comprises one or more secondary structure elements, amino acid residues, main chain atom groups, side chain atom groups, or individual atoms of the target molecule.
10. The method of claim 3, wherein each selected position comprises one or more bases, functional groups, or individual atoms of the target molecule.
11. The method of claim 1, wherein the interaction types represent different types of intermolecular interactions between the target molecule and the ligand.
12. The method of claim 1, wherein the interaction types is selected from the group consisting of contact interaction, polar interaction, non-polar interaction, or hydrogen bonding interaction.
13. The method of claim 11, wherein the intermolecular interactions are characterized by interaction energy-based approach.
14. The method of claim 12, wherein the contact interaction comprises an inter-atomic contact distance between the target molecule and the ligand of less than 10 Å.
15. The method of claim 12, wherein the contact interaction comprises an inter-atomic contact distance between the target molecule and the ligand of less than 6 Å.
16. The method of claim 12, wherein the contact interaction comprises a change in the accessible surface area of the target molecule upon forming a complex with the ligand.
17. The method of claim 12, wherein the hydrogen bonding interaction comprises a hydrogen bond donor in the target molecule and a hydrogen bond acceptor in the ligand at the selected position.

18. The method of claim 12, wherein the hydrogen bonding interaction comprises a hydrogen bond acceptor in the target molecule and a hydrogen bond donor in the ligand at the selected position.
19. The method of claim 1, wherein at least one interaction type includes a chemical or physical property of a part of ligand interacting with each selected position.
20. The method of claim 1, wherein each interaction type includes a chemical and physical property of a part of ligand interacting with each selected position.
21. The method of claim 19, wherein interaction types include information bits representing a core of a combinatorial library.
22. The method of claim 19, wherein interaction types include information bits representing a varying group of a combinatorial library.
23. The method of claim 19, wherein the interaction type includes a chemical property selected from the group consisting of hydrogen bond acceptor, hydrogen bond donor, hydrophobic, hydrophobic aliphatic, hydrophobic aromatic, negative charge, negative ionizable, positive charge, positive ionizable, or aromatic ring.
24. The method of claim 19, wherein the interaction type is an experimentally determined or computed property of the part of the ligand interacting with the selected position.
25. The method of claim 19, wherein the interaction type includes a chemical fingerprint for a part of the ligand interacting with the selected position of the target molecule.
26. The method of claim 1, wherein at least one interaction type includes a variable measuring the sequence conservation, structural conservation and flexibility of the selected position of the target molecule.
27. The method of claim 1, wherein the first ligand is the natural ligand of the target molecule or a ligand of known affinity to the target molecule.

28. The method of claim 1, further comprising calculating a score for each interaction among the target molecule-ligand complexes.
29. The method of claim 1, further comprising compiling the SIFts to generate an interaction profile from the calculated scores, wherein the interaction profile is indicative of the degrees of similarity of each information bit within the SIFts.
30. The method of claim 1, further comprising comparing a SIFt generated from a test ligand with an interaction profile generated from a group of target molecule-ligand complexes, thereby predicting whether the test ligand interacts with the target molecule in a similar pattern with the group.
31. The method of claim 1, further comprising comparing two interaction profiles, thereby predicting whether two groups of structures share conserved binding interactions, and/or have similar binding pattern.
32. A method of predicting the interaction pattern between a target molecule and a test ligand, the method comprising:
- identifying a plurality of selected positions between the target molecule and a first ligand, wherein the first ligand is known to bind to the target molecule;
 - generating a first structural interaction fingerprint (SIFt);
 - generating a second SIFt between the target molecule and a second ligand (test ligand) by repeating the identifying step and the generating a first SIFt step using the second ligand; and
 - comparing the first SIFt with the second SIFt to determine a level of overlapping, wherein a level of substantial overlapping predicts that each of the first ligand and the second ligand interacts with the target molecule in a similar pattern.
33. A method of generating a structural interaction fingerprint (SIFt) database, the method comprising:
- identifying a plurality of selected positions on a target molecule which is forming a complex with a first ligand;
 - generating a first SIFt of the database;

repeating the identifying step and the generating a first SIFt step using the same target molecule and a different ligand to generate a further SIFt of the database; and repeating the repeating step until the database contains a desired number of SIFts.

34. The method of claim 33, further comprising analyzing the SIFts of the database to generate one or more interaction patterns between the target molecule and the ligands.
35. The method of claim 34, further comprising comparing one interaction pattern of the database with a SIFt generated by using the same target molecule and a test ligand, thereby predicting whether said test ligand belongs to the family of ligands used to generate the database.
36. The method of claim 33, further comprising storing the database in a computer readable medium.
37. A method of analyzing the interaction pattern of two or more related target molecules, said method comprising:
- conducting sequence and structural alignments among each of the related target molecules to derive a uniform residue or base numbering system;
 - identifying a plurality of selected positions on the target molecule of each target molecule-ligand complex using one of the uniform residue or base numbering system;
 - generating a SIFt for each target molecule-ligand complex; and
 - comparing different SIFt patterns.
38. The method of claim 37, wherein at least one of the selected interactions among the complexes is conserved.
39. The method of claim 37, wherein at least one of the selected interactions among the complexes is unconserved.
40. The method of claim 37, wherein the related target molecules have at least 20% sequence similarity or a structural similarity with a root-mean squared deviation over the aligned positions no greater than 6 Å.

41. The method of claim 37, wherein the related target molecules have at least 20% sequence identity or a structural similarity with a root-mean squared deviation over the aligned positions no greater than 4 Å root-mean squared deviation over the aligned positions.
42. A computer-readable data storage medium comprising a data storage material encoded with a computer-readable database, the database comprising a plurality of structural interaction fingerprints (SIFts) generated from a target molecule and a plurality of ligands, wherein each SIFt is in the form of an information string comprising a plurality of information blocks and each information block comprising a plurality of information units, wherein said target molecule interacts with each ligand at a plurality of selected positions on the target molecule using a number of interaction types, and wherein the magnitude of each interaction type at each selected position is calculated and represented by a value which is assigned to an information unit.
43. The computer-readable data storage medium of claim 42, wherein the target molecule is a protein or a peptide.
44. The computer-readable data storage medium of claim 42, wherein the target molecule is a nucleic acid.
45. The computer-readable data storage medium of claim 42, wherein the ligand is a small molecule, a peptide, or a nucleic acid.
46. The computer-readable data storage medium of claim 42, wherein the value that is assigned to an information unit is a binary value, which indicates the presence or absence of a particular interaction type at the corresponding selected position.
47. The computer-readable data storage medium of claim 42, wherein the value that is assigned to an information unit is selected from a range of scaled numeric values, which indicates the magnitude of a particular interaction type at the corresponding selected position.

48. The computer-readable data storage medium of claim 42, wherein each selected position comprises one or more amino acid residues, main chain atom groups, side chain atom groups, or individual atoms of the target molecule.
49. The computer-readable data storage medium of claim 42, wherein each selected position comprises one or more bases, functional groups, or individual atoms of the target molecule.
50. The computer-readable data storage medium of claim 42, wherein each interaction type is selected from the group consisting of contact interaction, polar interaction, non-polar interaction, and hydrogen bond interaction.
51. The computer-readable data storage medium of claim 50, wherein the contact interaction comprises an inter-atomic contact distance between the target molecule and the ligand of less than 10 Å.
52. The computer-readable data storage medium of claim 50, wherein the contact interaction comprises an inter-atomic contact distance between the target molecule and the ligand of less than 6 Å.
53. The computer-readable data storage medium of claim 50, wherein the contact interaction comprises a change in the accessible surface area of the target molecule upon forming a complex with the ligand.
54. The computer-readable data storage medium of claim 50, wherein the hydrogen bond interaction comprises a hydrogen bond donor in the target molecule and a hydrogen bond acceptor in the ligand at the corresponding selected position.
55. The computer-readable data storage medium of claim 50, wherein the hydrogen bond interaction comprises a hydrogen bond acceptor in the target molecule and a hydrogen bond donor in the ligand at the corresponding selected position.
56. A computer program for generating a structural interaction fingerprint (SIFt) in the form of an information string which comprises a plurality of information blocks,

wherein each information block comprises a plurality of information units, the computer program comprising instructions for causing a computer system to:

select a plurality of selected positions on a target molecule wherein each selected position corresponds to an information block in the information string, the target molecule forming a complex with a ligand;

select a plurality of interaction types and calculate a value that is indicative of the characteristic of each interaction type at each selected position of the target molecule;

assign the value to a corresponding information unit, the information unit indicating a characteristic of the interaction type at the corresponding selected position;

join the information units of each selected position together to form corresponding information blocks; and

join the information blocks to generate a SIFT.

57. The computer program of claim 56, wherein the target molecule is a protein or a peptide.
58. The computer program of claim 56, wherein the target molecule is a nucleic acid.
59. The computer program of claim 56, wherein the ligand is a small molecule, a peptide, or a nucleic acid.
60. The computer program of claim 56, wherein the value that is assigned to an information unit is a binary value indicating the presence or absence of a particular interaction type at the corresponding selected position.
61. The computer program of claim 56, wherein the value that is assigned to an information unit is a numeric value selected from a scale of numbers, wherein the numeric value indicates the magnitude of a particular interaction type at the corresponding selected position.
62. The computer program of claim 56, wherein the selected positions are obtained from a three-dimensional structure of a binary complex formed between the target molecule and the ligand.

63. The computer program of claim 62, wherein the three-dimensional structure is derived from an experimental method or a prediction method.
64. The computer program of claim 57, wherein each selected position comprises one or more secondary structure elements, amino acid residues, main chain atom groups, side chain atom groups, or individual atoms of the target molecule.
65. The computer program of claim 58, wherein each selected position comprises one or more bases, functional groups, or individual atoms of the target molecule.
66. The computer program of claim 56, wherein the interaction types represent different types of intermolecular interactions between the target molecule and the ligand.
67. The computer program of claim 56, wherein the interaction types is selected from the group consisting of contact interaction, polar interaction, non-polar interaction, or hydrogen bonding interaction.
68. The computer program of claim 66, wherein the intermolecular interactions are characterized by interaction energy-based approach.
69. The computer program of claim 67, wherein the contact interaction comprises an inter-atomic contact distance between the target molecule and the ligand of less than 10 Å.
70. The computer program of claim 67, wherein the contact interaction comprises an inter-atomic contact distance between the target molecule and the ligand of less than 6 Å.
71. The computer program of claim 67, wherein the contact interaction comprises a change in the accessible surface area of the target molecule upon forming a complex with the ligand.

72. The computer program of claim 67, wherein the hydrogen bonding interaction comprises a hydrogen bond donor in the target molecule and a hydrogen bond acceptor in the ligand at the selected position.
73. The computer program of claim 56, further comprising instructions to store the SIFt in a database.
74. The computer program of claim 67, further comprising instructions to store the SIFts in a database.
75. The computer program of claim 56, further comprising instructions to generate a further SIFt between the target molecule and a second ligand without changing the selected positions and the interaction types; and comparing the original SIFt with the further SIFt to determine a level of overlapping, wherein a pattern of substantial overlapping predicts that each of the original ligand and the second ligand interacts with the target molecule in a similar pattern.
76. The computer program of claim 75, wherein the original ligand is the natural ligand of the target molecule.